

Fetal Health Classification

Présentation de l'étude

Contexte :

- Chaque année:
- 290 000 décès maternels
 - 1.9 million de morti-naissances
 - 2,3 millions de décès de nouveau-nés

Le **cardiotocogramme**, un examen qui enregistre la fréquence cardiaque fœtale et l'activité utérine de la mère.
Une meilleure interprétation → une meilleure prise en charge.

Jeu de données :

22 variables qui décrivent **2126 cardiotocogrammes**. Ces variables décrivent :

- l'histogramme des rythmes cardiaques du fœtus
- l'activité utérine de la mère (contractions, mouvements du fœtus...)
- Chaque individu a été classé parmi trois catégories : **Normal, Suspect, Pathologique**

Objectif : Prédire à quelle classe un individu appartient, tout en limitant les erreurs.

Exploration des données

- Variables prédictives : **21 variables quantitatives**
- Variable à prédire : **fetal health** qui est qualitative ordinale:
1 (Normal) < 2 (Suspect) < 3 (Pathological)

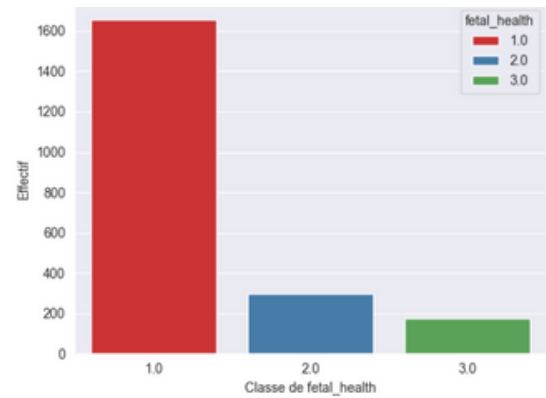


FIGURE 1 – Proportion des classes de la variable à prédire.

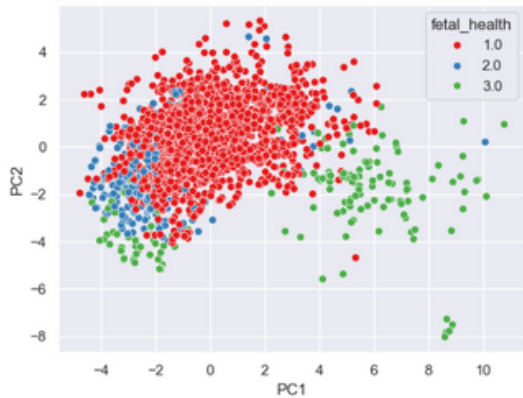


FIGURE 2 – Représentation des données dans le premier plan factoriel (ACP normalisée et pondérée)

Classification supervisée

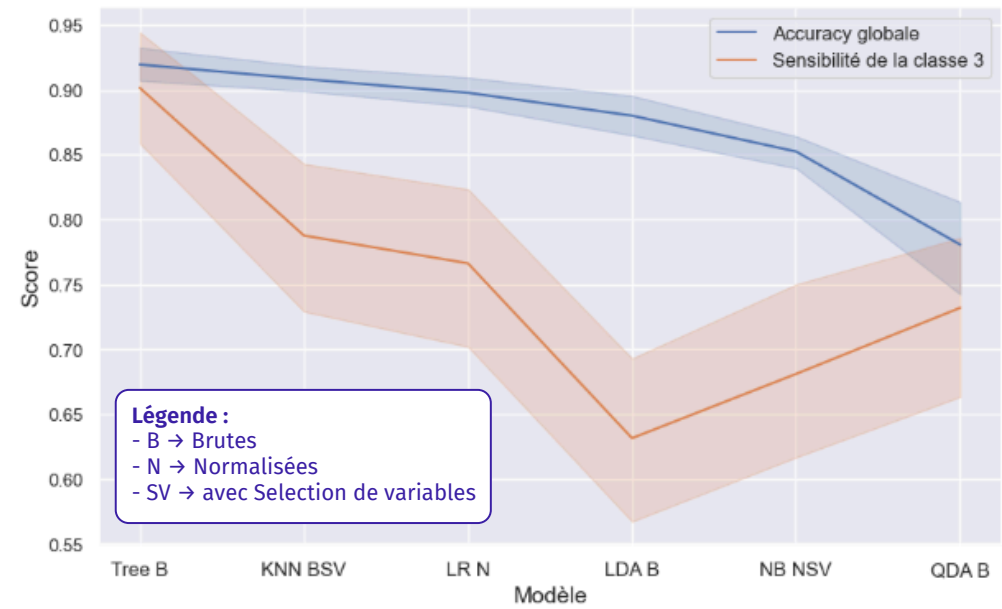


FIGURE 3 – Performances des différents classifieurs

Nos meilleurs classifieurs

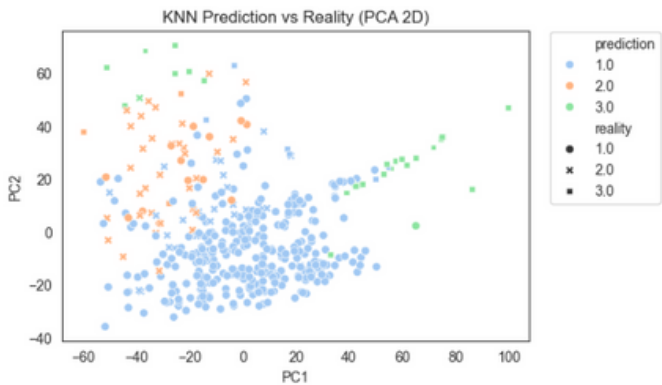


FIGURE 4 – Prédiction KNN VS Réalité (représentation dans le premier plan factoriel)

K Plus Proches Voisins :

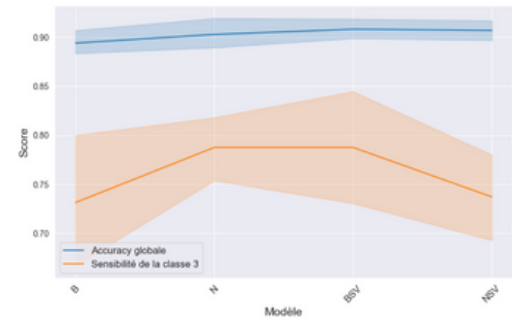


FIGURE 5 – Performances des 4 classifieurs KNN

TABLE 2 – Performances du classifieur KNN avec sélection de variables (moyennes des 10 plis de notre validation croisée)

Classe	Précision	Rappel	F1-score
Normal	0.94	0.96	0.95
Suspect	0.68	0.68	0.67
Pathological	0.91	0.79	0.84
Global	0.91		

Régression logistique :

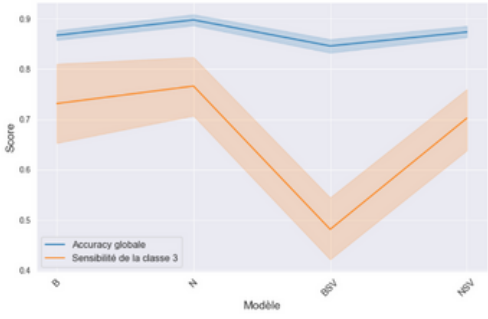


FIGURE 6 – Performances des 4 modèles de Régression Logistique

TABLE 3 – Performances de la régression logistique avec normalisation des données (moyennes des 10 plis de notre validation croisée)

Classe	Précision	Rappel	F1-score
Normal	0.95	0.95	0.95
Suspect	0.67	0.67	0.66
Pathological	0.85	0.77	0.80
Global	0.90		

Arbres de décision :

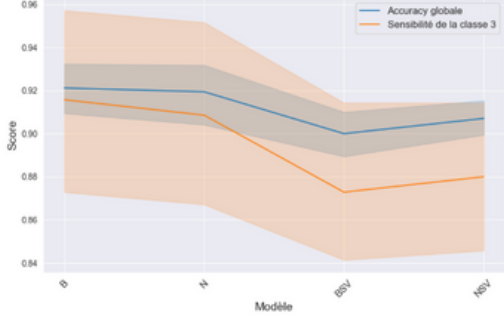


FIGURE 7 – Performances des 4 modèles d'Arbre de Décision

TABLE 4 – Performances de l'arbre de décision avec les données brutes (moyennes des 10 plis de notre validation croisée)

Classe	Précision	Rappel	F1-score
Normal	0.96	0.95	0.96
Suspect	0.75	0.76	0.76
Pathological	0.91	0.91	0.91
Global	0.92		